ILLUMINATION SYSTEM

Docket Number: 200400376-1

Inventor:

Norman Paul Jouppi
Citizenship: USA
617 Colorado Ave.
Palo Alto, CA 94306

ILLUMINATION SYSTEM

BACKGROUND

TECHNICAL FIELD

The present invention relates to an illumination system.

BACKGROUND ART

In robotic telepresence, a remotely controlled robot simulates the presence of a user. The overall experience for the user and the participants interacting with the robotic telepresence device is similar to videoconferencing, except that the user has a freedom of motion and control over the robot and video input that is not present in videoconferencing. The robot platform or surrogate typically includes a camera, a display device, a motorized platform that includes batteries, a control computer, and a wireless computer network connection. An image of the user is captured by cameras at the user's location and displayed on the robotic telepresence device's display in the surrogate's location.

DISCLOSURE OF THE INVENTION

The present invention provides an illumination control system including viewing illumination at a surrogate's location and recreating the illumination at a user's location as a relative perceived illumination.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGs. 1A and 1B show an embodiment of a Mutually-Immersive Mobile Telepresence (E-Travel) System;

FIG. 2 shows a luminance diagram illustrating four cameras in the surrogate's head pointing in different outward directions;

FIG. 3 shows an embodiment of a camera on the surrogate's head and a light sensor assembly;

FIG. 4 shows a luminance diagram similar to FIG. 2 where the luminance is scaled linearly from a midpoint by a computer;

FIG. 5 shows a luminance diagram similar to FIG. 2 where the luminance is scaled linearly from the brightest video by a computer;

FIG. 6 shows a luminance diagram similar to FIG. 2 where the luminance is scaled non-linearly from a midpoint by a computer;

FIG. 7 shows a luminance diagram similar to FIG. 2 where the luminance is scaled with varying base illumination by a computer; and

FIG. 8 shows a system 800 of mobile telepresencing according to an embodiment of the present invention.

BEST MODE FOR CARRYING OUT THE INVENTION

A Mutually-Immersive Mobile Telepresence System may have a robot device of a humanoid as well as a non-humanoid shape, which is referred to as a "surrogate". A user sits in a room that may show the surrogate's location and the surrogate may be located at a surrogate's location. Video and audio are transmitted between a user display and the surrogate. The user sees views radially outward on the user display and 360 degrees around from the center of the surrogate to have the feeling of being present at the surrogate's location by seeing it in a surround view, and the people or meeting participants at the surrogate's location have the feeling that the user is present by display panels on the surrogate showing images of the head of the user; i.e., the feeling of telepresence.

The user sits or stands inside a display cube, with rear-projection surfaces on the front, back, sides, and optionally the ceiling showing the surrogate's location. Since the goal is to be mutually immersive, live color video images of the user centered on the user's head are acquired from all four sides of the user's location for transmission to the surrogate's location concurrent with projection of live color video surround from the surrogate's location on the four sides of the display cube surrounding the user. The user can move about inside the display cube, so head tracking techniques are used to acquire pleasingly cropped color video images of the user's head in real time.

Referring now to FIGs. 1A and 1B, therein are shown a Mutually-Immersive Mobile Telepresence System 100, which includes a display cube 101 at a user's location 104 and a surrogate 106 at a surrogate's location 108. The surrogate 106 is connected to the display cube 101 via a high-speed network 110.

The surrogate 106 has a surrogate's head 112 including a number of head display panels 114, such as four LCD panels. Video equipment in the form of one or more cameras, such as four surrogate's cameras 116-1 through 4, is positioned in the corners of the surrogate's head 112 to view and capture 360 degrees surround live video at the surrogate's location 108 for display on the display cube 101. An optional outward facing light sensor 117, facing up in FIG. 1A, may also be positioned on the top of surrogate's head 112 to detect the brightness of light overhead.

One or more microphones, such as four directional surrogate's microphones 118, are positioned in the top corners of the surrogate's head 112 to capture sounds 360 degrees around the surrogate 106. One or more speakers, such as the four surrogate's speakers 120 are also positioned in the bottom corners of the surrogate's head 112 to provide directional audio of the user's voice.

The surrogate 106 contains surrogate's computer/transceiver system 122 connecting the surrogate's cameras 116-1 through 4, the surrogate's microphones 118, and the surrogate's speakers 120 with the display cube 101 for a user 124. The surrogate's computer/transceiver system 122 also receive live video views of the user's head 126 from user's cameras 128-1 through 4 at the four corners of the display cube 101 and display the live video views on the head display panels 114 in the surrogate's head 112.

The display cube 101 at the user's location 104 receives the video and audio signals at a user's computer/transceiver system 130. Video equipment provides views from the surrogate's cameras, such as the surrogate's cameras 116-1 through 4 in the surrogate's head 112, which are projected on projection screens 102 of the display cube 101 by projectors, such as four user's projectors 132-1 through 4. An optional variable brightness projector, such as a dimmable light 133, may be positioned facing inward, or down from above the user's head 126 in FIG. 1A.

User's speakers 134 are mounted above and below each projection screen 102. By driving each pair of user's speakers 134 with equal volume signals the sound appears to come from the center of each of the projection screens 102 to provide directional audio or hearing of one or more participants from the surrogate's microphones 118.

The user's computer/transceiver system 130, which can be placed in an adjacent room (for sound isolation purposes), drive the user's speakers 134 with audio information transmitted from the surrogate 106 at the surrogate's location 108.

The images on the projection screens 102 are presented "life size". This means that the angle subtended by objects on the projection screens 102 is roughly the same angle as if the user 124 was actually at the surrogate's location 108 viewing it personally when the user's head is centered in the display cube 101.

To have full surrogate mobility, the surrogate 106 has remote translation and remote rotation capabilities. The term "translation" herein means linear movement of the surrogate 106 and the term "rotation" herein means turning movement of the surrogate 106.

When the user 124 desires to change body orientation with respect to the surrogate's location 108, the user 124 may do so by turning at the user's location 104 and having the surrogate 106 remain stationary but the head display panels 114 on the surrogate 106 show the user's head 126 turning to face the desired direction without movement or a rotation of the surrogate 106.

The surrogate 106 has a surrogate's body 140, which is rotationally (circularly) symmetric and has no front, back or sides (i.e., the base and body of the surrogate 106 are cylindrical). Furthermore, the surrogate 106 uses a mechanical drive system 144 that can travel in any translational direction without a need for rotation of the surrogate's body 140.

As part of mutually immersive telepresence, the user 124 in the display cube 101 should be lit as if they were at the surrogate's location 108. Thus, the system 100 preserves the relative illumination levels of the surrogate's location 108. For example, if the surrogate 106 is in an office with the user's face image facing a window during daylight hours, the front of the user's face should be brighter than the back of the user's head. This would produce the same effect for the user 124 and the participants at the surrogate's location 108 as if the user 124 was physically present at the surrogate's location 108. In order to achieve this effect, the user 124 must first see the window area as brighter than the office area behind the user's back when the user 124 is in the display cube 101.

To further enhance the effect of mutually immersive telepresence, the light sensor 117 would control the dimmable light 133 to provide the same overhead lighting effect produced by overhead lighting at the surrogates's location 108. It is also possible to apply the light sensor and variable light combination in general for any direction where it is decided that a projector and screen combination is not desired or is not necessary.

It has been discovered that, due to adaptation of the human visual system (described below), it is not necessary to recreate the absolute luminance values of the surrogate's location 108. To do so would be expensive, since even the brightest projectors project at much dimmer light levels than daylight light level scenes. Instead, it has been discovered that it is possible to recreate the relative perceived luminance of different portions of the surrogate's location 108 to preserve proper illumination.

The human visual system can perceive different scenes that vary in brightness by a factor of more than 10 million. Human vision can adjust to varying illumination levels by the process of adaptation. Humans adapt both by changing the open portion of their pupil from roughly 2 mm to 7 mm in diameter as well as changing the response characteristics of their rods and cones via both photochemical and neural processes. Cameras can also image scenes that vary widely in illumination, and adapt their exposure settings to compensate for illumination. The principal camera adaptations are shutter speed and the variation of a mechanical iris opening.

Adaptation is required to adjust to very large changes in illumination (e.g., of a million fold).

The range of simultaneous human illumination perception is smaller, and is around 13 camera f-stops. In other words, humans can see details in both highlights and shadows in scenes that vary in illumination from 2^{13}, or roughly 8192 times, between the highlight and shadow detail. However, current commodity cameras generally have a maximum capture range of only 8 bits, or 2^{8}. This gives such cameras a dynamic range of 256 times, which is 32 times smaller than human vision.

It has been found that useful telepresence systems will invariably need to employ multiple cameras for the foreseeable future. This is because the resolution of a single camera is so far below that of human visual acuity. Thus, multiple cameras must be used to increase the effective resolution, and limitations of optics limit the resolution of a single camera lens. Human foveal visual resolution mapped onto a 360 degree surround horizontal field of view and a 60 degree vertical field of view would be equivalent to a 44,000 by 7,300 pixel display.

It has also been found that, when multiple cameras are used, each one will typically make its own adaptation to best fit its portion of the scene into its own restricted simultaneous dynamic range (e.g., 8 bits). This means that each camera typically will use its own exposure settings. If a panorama with different exposure settings is naively joined together, discontinuities will result at the borders between images. For example, a camera facing the

windows would have a faster shutter speed than a camera facing the interior of the room, so its image would appear darker at the boundary where the corresponding image facing the interior begins.

Referring now to FIG. 2, therein is shown a luminance diagram 200 having a scale of 0 to 8 log candelas per square meter (cd/m^2) illustrating four cameras 116-1 through 4 in the surrogate's head 112 pointing in different outward directions, and each having a different exposure setting 201 through 204 based on the illumination that it sees. This results in each camera's 8-bit video data corresponding to different illumination ranges. In effect, the actual video values are similar to the mantissa in scientific notation, while the exponents in traditional video are discarded. (For example, 2.345×10^6 is an example of scientific notation, with 2.345 being the mantissa.) Hence, a standard video stream only records relative luminance information 205, and not absolute luminance information.

Besides being limited on camera acquisition, dynamic range is also limited with modern projector technology for the display cube 101. Unlike with cameras, the illumination of the projectors 132-1 through 4 is not typically adaptable. Furthermore, the simultaneous dynamic range is likewise typically limited to a range of 256:1, or 8-bit precision. Thus, given a matched set of projectors 132-1 through 4 it is not possible to project imagery that varies by more than a 256:1 ratio representing a bright side to a dark side of the surrogate's location 108.

It was discovered that a system that recorded the absolute luminance in each camera's view of the surrogate's location 108, and transferred this record to the user's location 104 creates a better display of the surrogate's location 108. The one way of obtaining the absolute luminance value for each camera 116-1 through 116-4 viewing the surrogate's location 108 is to use a digital video camera that could read back the exposure settings directly. However, these cameras are currently expensive and the video compression hardware interfacing to the camera would also need to support access to this data. Conventional video cameras do not output exposure information, only analog composite video or S-video signals.

Referring now to FIG. 3, therein is shown one of the cameras 116 on the surrogate's head 112, such as the camera 116-1, having a light sensor assembly 300, such as the light sensor assembly 300-1, mounted just above the camera 116-1. The light sensor assembly 300-1 provides an estimate of the exposure settings of the camera 116-1 to a microcontroller (not shown) and that in turn interfaces to the surrogate's computer/transceiver system 122.

The light sensor assembly 300-1 has two light sensors 302 and 304 for accurately capturing a very wide range of luminance values at relatively low cost. To better approximate the visual sensitivity of the human eye, the two light sensors are behind a blue-green filter 306. One sensor 302 is configured to accurately measure relatively bright environments such as outdoor scenes, and one sensor 304 is configured to measure relatively dim environments such as conference rooms. The light sensor assembly 300-1 is placed in a housing such that it has roughly the same field of view as the lens on the camera 116-1. This is done to minimize the differences in the scenes viewed by the camera 116-1 and the light sensors 302 and 304. In an embodiment with tilting cameras, the light sensors 302 and 304 could also be tilted with the camera 116.

Knowing the absolute luminance seen by each camera 116-1 through 4, the relative exposure settings can be estimated. This estimation is more accurate if the camera exposure is controlled in a relatively straightforward fashion, without compensating for backlighting. All the luminance values are provided to the user's computer/transceiver system 130. Once the user's computer/transceiver system 130 knows the relative average luminance of each video stream, many different approaches are possible.

Referring now to FIG. 4, therein is shown a luminance diagram similar to FIG. 2 where the luminance is scaled linearly from a midpoint by a computer, such as the user's computer/transceiver system 130. The user's computer/transceiver system 130 computes the average luminance seen in each video stream, and scale each video stream based on the ratio of its luminance to the average log of the luminance.

For example, if two video streams had equal luminance L, one stream's luminance was twice the average, and one stream's luminance was half the average, the average of the log of the luminance would be log L. Then, all the video values in the stream pointing in the 2X brighter direction, such as would have their color values multiplied by 2X, while the stream that was 2X dimmer would have their color values multiplied by 1/2X.

Pixels that had colors that overflowed after multiplication would be set to the maximum brightness possible for that hue. For example, a color described by RGB = [100/255, 0/255, 200/255] could only be increased to RGB = [128/255, 0/255, 255/255] without significantly changing the hue of the pixel. (Changing the hue of the pixel would change the color of the lighting of the user, and so is undesirable.) Hue changes are also a function of gamma, so the example above also assumes a gamma equal to one. However, to the extent that some pixels will be able to scale fully, and other pixels will only be able to

undergo limited scaling, artifacts in the video images will be introduced. In contrast, pixels that underflowed to zero as a result of being multiplied by factors of less than one could simply be set to zero.

FIG. 4 shows a luminance diagram 400 having a scale of 0 to 8 log candelas per square meter (cd/m^2) illustrating illumination from the four projectors 132-1 through 4 in the display cube 101 of FIG. 1A. The four projectors 132-1 through 4 provide the projector levels 401 through 404 using linear scaling based on a midpoint for the surrogate's luminance range data to provide a full projection range 405. The projector levels 401 and 403 are at the full projection range 405 and the projector levels 402 and 404 are the same but one is at the lower end of the full projection range 405 and the other is at the upper end. It should be noted that the absolute illumination levels available from even high-quality projectors are significantly less than the outdoor scenery illumination levels of the first figure. However, the absolute illumination levels available with projectors can be close to that of office environments.

Referring now to FIG. 5, therein is shown a luminance diagram similar to FIG. 2 where the luminance is scaled linearly from the brightest video by a computer, such as the user's computer/transceiver system 130 of FIG. 1A. From the above discussion regarding overflows, it is clear that pixel values cannot be increased without causing color shifts, image distortions, and saturation. Thus, a more accurate approach would be for the user's computer/transceiver system 130 to define the video stream corresponding to the brightest direction to have a scaling factor of one. Then, all of the darker video streams would need to have their pixel colors multiplied by a scaling factor that was less than one. This would prevent the saturation artifacts caused by increasing pixel values, but could lead to a large number of pixels in other streams becoming black.

Considering the example above again, the brightest stream would remain unchanged, two streams would have their pixel luminance multiplied by ½ and the darkest side would have its luminance multiplied by ¼. Multiplying pixel values by ¼ is equivalent to converting from an 8-bit to a 6-bit color. If less than 8 bits are used to describe pixel colors, banding artifacts can become visible. In situations with even greater illumination differences between video streams, dark streams could be reduced to 4-bit color or even turn completely black in the most extreme case.

FIG. 5 shows a luminance diagram 500 having a scale of 0 to 8 log candelas per square meter (cd/m^2) illustrating illumination from the four projectors 132-1 through 4 in the

display cube 101 of FIG. 1A. The four projectors 132-1 through 4 provide the projector levels 501 through 504 using linear scaling based on a midpoint for the surrogate's luminance range data to provide a full projection range 505. The projector level 504 is at the full projection range 505, the projector levels 501 and 503 are the same at three-quarters of the full projection range 505, and the projector level 502 is at a half of the full projection range 505.

There are several problems with the above approach. First, if extreme modifications to lighting the user 124 are made, it can make acquisition of the user's image difficult. The cameras 128 of FIG. 1A on the dark side of the user 124 will experience strong backlighting, and image detail in the user's body will be lost. This is one consequence of video cameras having a smaller simultaneous dynamic range than human vision. Another problem occurs because projectors also have a smaller simultaneous dynamic range than with human vision. (Actually it is even worse, since a user's vision will adapt when looking only at two adjacent dark sides of a display cube while the projector cannot.)

While this would definitely be a more accurate recreation of the illumination levels at the surrogate's location 108, such extreme changes may be objectionable for the user 124. Instead, more moderate adjustments to illumination levels could prove more attractive.

Referring now to FIG. 6, therein is shown a luminance diagram similar to FIG. 2 where the luminance is scaled non-linearly from a midpoint by a computer, such as the user's computer/transceiver system 130 of FIG. 1A. The user's computer/transceiver system 130 computes a non-linear function of the luminance difference, such as a square root.

For example, a 4X luminance difference would result in the luminance of the darker video only being divided by two. This has the advantage that distinctions would still be made between video streams at different illumination levels, but problems due to limited camera and projector simultaneous dynamic range would be reduced. It should be noted that since gamma characterization of cameras and projectors is not exact, even with gamma=1 the display characteristics would not be precisely linear, so intensity gaps and overlaps could occur with linear mapping.

Furthermore, during back projection even with a screen having a nominal gain of unity, images will tend to be brighter at their centers than around their edges. Such brightness variations are a function of the user's position relative to the various screens. Thus exact matching of luminance levels between screens is not necessary given the magnitude of other

luminance errors. An example showing square root scaling from the midpoint is shown below. Note that each projected view retains a wider simultaneous dynamic range.

FIG. 6 shows a luminance diagram 600 having a scale of 0 to 8 log candelas per square meter (cd/m^2) illustrating illumination from the four projectors 132-1 through 4 in the display cube 101 of FIG. 1A. The four projectors 132-1 through 4 provide the projector levels 601 through 604 using non-linear, square root scaling based on a midpoint for the surrogate's luminance range data to provide a full projection range 605. The projector levels 601 and 603 are at the full projection range 605 and the projector levels 602 and 604 are the same but one is a square root at the lower end of the full projection range 605 and the other is a square root at the upper end.

It has been discovered that a user adjustable scaling reference is a useful option. This allows the user 124 to select between the midpoint, brightest video, darkest video, or even points in between as a scaling baseline. One user interface for this could be the thrust wheel of a joystick, with the middle wheel rotation denoting scaling from the midpoint. Having the ability to adjust the scaling reference would allow the user 124 to recover details lost in one video stream either due to saturation or a reduction in the effective number of bits.

It has also been discovered that ramping luminance at screen boundaries is another useful option. If a non-linear scaling is performed, the edges of adjacent video streams will still exhibit luminance discontinuities (i.e., one will be darker than another). One way to adjust this would be to use the user's computer/transceiver system 130 of FIG. 1A to modify the luminance scaling in the horizontal edges of the video streams to locally improve the matching. If non-linear scaling is performed, the video streams that should be dark will be brighter than they would be with linear scaling. Hence, as they approach an adjacent brighter video stream they should have their luminance ramp down, while if they approach a dimmer video stream they should have their luminance ramp up. This ramping is probably best done via a smooth function such as a sine function, so that both the ramping and the derivative of the ramping are continuous.

Referring now to FIG. 7, therein is shown a luminance diagram similar to FIG. 2 where the luminance is scaled with varying base illumination by a computer, such as the user's computer/transceiver system 130 of FIG. 1A. The user's computer/transceiver system 130 computes the baseline illumination level of each projector 132 based on the luminance recorded by the corresponding light sensor 300 on the surrogate 106.

The major problem with the methods discussed so far is that the simultaneous dynamic range of projectors are less than the simultaneous dynamic range of human vision, and that restricting the projector range further to better match luminance levels can degrade image quality significantly.

It has been discovered that varying the baseline illumination level of each projector based on the luminance recorded by the corresponding light sensor allows each projector 132 to retain its simultaneous dynamic range, while at the same time linearly matching illumination levels.

FIG. 7 shows potential projection ranges using this technique, assuming that the projector light output can be reduced by up to 10X. A luminance diagram 700 is shown having a scale of 0 to 8 log candelas per square meter (cd/m^2) illustrating illumination from the four projectors 132-1 through 4 in the display cube 101 of FIG. 1A. The four projectors 132-1 through 4 provide the projector levels 701 through 704 by varying the baseline illumination level so the luminance starts at different points. The projector levels 701 through 704 have linearly matching illumination levels. As a result, the projector levels have a total projector range 705.

There are many possible methods for varying baseline illumination levels of the projectors 132. The equipment can be ancillary to or in the projectors 132.

A first method would be to vary the voltage going to the projector's lamp. However, this can change the color temperature output by the bulb, and would need to be compensated for by other means. Nevertheless, this is a method for conventional (mono) projection. Lamps should have their lifetimes increased by reducing their power levels, so the main cost of this technique would be the variable voltage lamp power supply and color correction.

A second method would be to use electrochromic glass, which can change its transmission based on an applied electric field. However, electrochromic glass currently does not support a wide dynamic range, has a color cast when it is less transmissive, is heat sensitive (anything blocking light from a 250W bulb will get hot), and has slow response times.

A third method would be to insert a first fixed polarizing filter and a second rotating polarizing filter on the output of the projector lens. The rotating polarizing filter could be rotated under computer control to vary the amount of light absorbed by the pair. As the second polarizing filter is rotated out of parallel with the first, the light retains the polarization of the first filter but gets progressive dimmer as a 90-degree offset is approached.

For conventional (mono) video projection this technique has the disadvantage that a pair of parallel polarizing filters transmit at most 25% of the original light field. (This corresponds to the 2 *f*-stop loss in photography for polarizing filters.) Thus, the absolute light levels in the display cube 101 would be significantly reduced, which would make acquisition of the user's image more difficult. However, this method has the advantage for stereo projection where projectors supporting 3D vision often already have polarized output. This would lead to negligible light loss in the case of an additional parallel polarizing filter. Thus, this technique is preferred for stereo projection applications.

Referring now to FIG. 8, therein is shown an embodiment of a method or system 800 of mobile telepresencing, including a block 802 of viewing illumination at a surrogate's location and a block 804 of recreating the illumination at a user's location as a relative perceived illumination.

While the invention has been described in conjunction with a specific best mode, it is to be understood that many alternatives, modifications, and variations will be apparent to those skilled in the art in light of the aforegoing description. Accordingly, it is intended to embrace all such alternatives, modifications, and variations which fall within the scope of the included claims. All matters hither-to-fore set forth herein or shown in the accompanying drawings are to be interpreted in an illustrative and non-limiting sense.